TITLE: Dependence Structure of Random Wavelet Coefficients in Terms of Cumulants

Dependence Structure of Random Wavelet Coefficients in Terms of Cumulants

Philippe Naveau, Peter Brockwell, and Doug Nychka

Abstract. When the Gaussian assumption for a times series no longer holds, second order moment properties such as the covariance and the spectrum are not necessarily sufficient to describe the dependence structure. Although wavelet models have been proposed to de-correlate the signal, this strategy must be reexamined when applied to non-Gaussian processes. The process of interest is a continuous parameter, mean-squared continuous real-valued process that is not necessarily Gaussian or linear. To study the departures from linearity and Gaussianity, we consider joint cumulants, which are linear combinations of higher order moments, and their associated spectra. A specific objective is to obtain new expressions for cumulants of the random discrete wavelet coefficients instead of the second order moments, and to study their higher order polyspectra. Conditions on the polyspectrum to give null wavelet cumulants within and across wavelet coefficient levels are derived. Expressions of the original cumulants as a function of the wavelets cumulants are also given.

§1. Introduction

The covariance and spectral properties of the discrete wavelet coefficients for random continuous real-valued processes have been extensively studied in the past. Among others, Donoho et al. [4], Flandrin [6], Mallat et al. [13], Masry [14], and Walter [17] have investigated the correlation within and across wavelet coefficients. Focusing exclusively on the second order properties of wavelet coefficients for a Gaussian process is a reasonable task since the dependence structure of Gaussian processes is entirely characterized by the covariance. When the normality assumption no longer holds, higher order cumulants are necessary.

Exploring some of the links that exist between wavelets and cumulants is fairly new. Brillinger [1] studied a non-parametric regression problem with cumulants and wavelets. In geophysics and astrophysics, Lazear [11] and Ferreira et al. [5] applied wavelets and cumulants to seismic data sets and to the Cosmic Microwave Background problem. The dependence structure

Curve and Surface Fitting: Saint-Malo 1999
Albert Cohen, Christophe Rabut, and Larry L. Schumaker (eds.), pp. 343–350.

ISBN 0-8265-1357-3.

between wavelet packets and in particular wavelet coefficients has also been studied by D. Leporini and J.C. Pesquet [12]. In this paper, we derive new and general results about the dependence structure of random wavelet coefficients via its cumulants.

The process of interest $X(t)$, indexed by the real parameter t, is supposed to be a continuous real-valued process that is mean-square continuous, and such that moments of some order $l \geq 2$ exist, i.e.,

$$\sup_t E|X(t)|^l < \infty \qquad \text{and} \qquad \lim_{h \to 0} E|X(t+h) - X(t)|^2 = 0. \tag{1}$$

§2. Cumulant Definition and Properties

If some useful information of the signal is not contained in the second-order covariances (and the second order spectra), then one can still calculate some meaningful linear combination of higher order moments, called cumulants. Some early work on higher order cumulants and their Fourier transform was proposed by Hasselman et al. [8] for investigating nonlinear interaction of ocean waves, and Godfrey [7] used it for the analysis of economic time series. Rosenblatt with Lii and Van Atta in a series of papers have described how higher cumulants could be used to study nonlinear transfer of energy in turbulence.

The m^{th} joint cumulant of the set of random variables $\{X(t_1), ..., X(t_m)\}$, denoted by $CUM(X(t_1), ..., X(t_m))$, with $m \leq l$, is given by

$$CUM(X(t_1), ..., X(t_m)) = \sum (-1)^p (p-1)! (E \prod_{r \in v_1} X(t_r)) ... (E \prod_{r \in v_p} X(t_r)),$$

where the summation extends over all partitions $\{v_1, ..., v_p\}$ of $\{1, ..., m\}$ with $p = 1, ..., m$. From this definition, we can notice that the information contained in the first m cumulants is exactly the same as that contained in the first m moments. However, cumulants have some advantages over moments. For example, cumulants have useful linear properties,

$$CUM(Z + X(t_1), ..., X(t_m)) = CUM(Z, ..., X(t_m)) + CUM(X(t_1), ..., X(t_m)),$$
$$CUM(aX(t_1), ..., X(t_m)) = aCUM(X(t_1), ..., X(t_m)),$$

for any real a. Another important property of cumulants concerns the dependence structure of the process: if some subset of $\{X(t_1), ..., X(t_m)\}$ is independent of the remainder, then $CUM(X(t_1), ..., X(t_m))$ is identically equal to zero. Hence, the cumulant, $CUM(X(t_1), ..., X(t_m))$ can be interpreted as a measure of dependence of $\{X(t_1), ..., X(t_m)\}$. For the special case of Gaussian processes, cumulants of order higher than two are zero.

In the remainder of this section, we suppose that the process $\{X(t)\}$ is stationary up to order l, i.e.,

$$E(X(t_0)X(t_1)...X(t_l)) = E(X(t_0+h)X(t_1+h)...X(t_l+h)), \qquad \forall h.$$

Stationarity as just defined is frequently referred to in the literature as weak stationarity. For us however the term stationarity, without further qualification, will always refer to the above equality.

Under stationarity, $CUM(X(t), X(t+s_1), ..., X(t+s_{m-1}))$ does not depend on t, and can be denoted by $\gamma_m(s_1, ..., s_{m-1})$. With these notations, the second order cumulant $\gamma_1(u)$ is just the covariance function. The third order cumulant $\gamma_2(u,v)$ is the same as the third-order central moment,

$$E((X(t)-\mu)(X(t+u)-\mu)(X(t+v)-\mu)),$$

where μ is the mean value of the process.

From the covariance function, one may define the power spectrum, i.e., the Fourier transform of $\gamma_1(t)$, $f_1(\omega) = \int \gamma_1(t)\exp(-i\omega t)dt$. A natural extension is the m^{th}-order polyspectrum defined by

$$f_m(\omega_1, ..., \omega_{m-1}) = \int ... \int \gamma_m(t_1, ..., t_{m-1}) \exp(-i \sum_{j=1}^{m-1} \omega_j t_j) dt_1 ... dt_{m-1},$$

assuming that the above Fourier transforms exist. An important property of the polyspectra is that all polyspectra of higher order than second order vanish when $\{X(t)\}$ is a Gaussian process. Another characteristic of the polyspectra is that the ratio

$$\frac{|f_2(\omega_1, \omega_2)|^2}{f_1(\omega_1) f_1(\omega_2) f_1(\omega_1 + \omega_2)}$$

is constant whenever the process $\{X(t)\}$ is linear. Hence, the simplest higher order spectrum, called bispectrum, can be regarded as deviation measures from Gaussianity and linearity. Different statistical tests have been derived from it (see Subbua Rao and Gabr [9] and Hinich [16]).

§3. Random Wavelet Coefficients

Consider a discrete orthonormal wavelet decomposition of a stochastic process $\{X(t)\}$ that satisfies condition (1). The corresponding wavelet coefficients

$$W_{j,k} = \int X(t) \psi_{j,k}(t) dt \tag{2}$$

are random variables. Here the equality sign is to be understood in the mean-square sense, and $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ is an orthonormal wavelet basis function with the mother wavelet ψ. A rigorous framework concerning the construction for wavelet orthonormal basis can be found in Meyer [15] and Daubechies [3]. There exist many candidates for the mother wavelet. The simplest example of an orthonormal wavelet basis is provided by the Haar system for which

$$\psi(x) = \begin{cases} 1, & \text{if } 0 \le x < 0.5, \\ -1, & \text{if } 0.5 \le x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Our main interest is to understand the dependence structure between the different wavelet coefficients defined by (2).

In this article, we will either use the Haar system as a simple example of compactly support wavelets, or a particular type of band-limited wavelets called Meyer-type wavelets (see Walter [17], Zayed and Walter [18]). Having a compact support in the frequency domain can facilitate the computation of wavelet cumulants. These Meyer-type wavelets have some additional attractive features, such as being highly smooth (they can be made C^∞) and having fast decay in the time domain. They are introduced as follows.

Let F be any probability measure supported on $[-\epsilon, \epsilon]$ for some $\epsilon \leq \pi/3$. Then the mother wavelet $\psi(\cdot)$ is defined by its Fourier transform

$$\tilde{\psi}(\omega) = \exp(-i\omega/2)[\int_{|\omega/2|-\pi}^{|\omega|-\pi} dF]^{1/2}. \tag{3}$$

From this definition, it is possible to check that the orthogonality and dilation conditions are satisfied for the wavelet basis generated from this mother wavelet. There is a large class of distributions that can be chosen in equality (3), and the edges of the support, $[-\epsilon, \epsilon]$, can be made highly smooth.

§4. Dependence Structure

The dependence structure between wavelet coefficients is closely related to the dependence inside the original signal. Hence, our first problem is to explain how to obtain the joint cumulants of the wavelet coefficients from the joint cumulants of the process. The first proposition takes care of this problem.

Because of space limitations, the proof of our propositions will not be included in this paper. However, complete details of the proofs can be requested from the authors.

Proposition 1. *Let $\{X(t)\}$ be a stochastic process that satisfies condition (1). Suppose that the joint cumulants of order $m \leq l$ of $\{X(t)\}$ exist. Then*

$$CUM(W_{j_1k_1}, ..., W_{j_mk_m}) = \int ... \int CUM(X(t_1), ..., X(t_m)) \prod_{n=1}^{m} \psi_{j_n,k_n}(t_n) dt_n.$$

Proposition 1 is directly applicable to compactly supported wavelets, since the product

$$\prod_{n=1}^{m} \psi_{j_n,k_n}(t_n)$$

is null except at the intersection of the translated and dilated supports. For example, suppose that the wavelet basis corresponds to the Haar system. The expression of cumulants of the wavelet coefficients becomes

$$CUM(W_{j_1k_1}, ..., W_{j_mk_m}) = \sum_{l=1}^{2^m} (-1)^l 2^{j_l/2} \int_{A_l} CUM(X(t_1), ..., X(t_m)) dt_n. \tag{4}$$

The A_l are rectangular boxes defined by the tensor product

$$A_l = \bigotimes_{n=1,..,m} I_{n,c_{nl}},$$

where the vectors $\tilde{c}_l = (c_{1l},, ..., c_{ml})$ represent the set of all 2^m possible configurations of $\{0, 1/2\}^m$, and the interval $I_{n,c_{nl}}$ is defined by

$$I_{n,c_{nl}} = \{t_n : 2^{-j_n}(k_n + c_{nl}) \leq t_m < 2^{-j_n}(k_n + c_{nl} + 1/2), \text{ for } n = 1, .., m\}.$$

In order to apply (4) to a more specific example, we suppose that the process $\{X(t)\}$ is a zero-mean stationary process with standard deviation σ and covariance $\gamma_1(h) = b\exp(-a|h|)$, and $\gamma_1(h_1, h_2) = c\exp(-a|h_1 + h_2|)$, where a, b, c are constants that depends on the second and third moment and other parameters describing the original process. Processes with such a cumulant function correspond to Continuous Auto-Regressive processes (CAR) (see Brockwell [2]) or equivalently solutions of particular stochastic differential equations with non-necessarly Gaussian noise. After some algebra, wavelet cumulants simplify to

$$CUM(W_{j_1k_1}, W_{j_2k_2}) = -\frac{bK_a(j_1, k_1)}{a^2K_a(j_2, k_2)}[\sum_{(u_1,u_2)\in\{-1,1\}^2} H_{j_1}(u_1a)H_{j_2}(u_2a)],$$

for $2^{-j_2}(k_2 + 1) \leq 2^{-j_1}k_1$ and

$$CUM(W_{j_1k_1}, W_{j_2k_2}, W_{j_3k_3}) = -\frac{cK_{2a}(j_1, k_1)}{2a^3K_a(j_2, k_2)K_a(j_3, k_3)} \times [\sum_{(u_1,u_2,u_3)\in\{-1,1\}^3} H_{j_1}(u_1a)H_{j_2}(u_2a)H_{j_3}(u_3a)]$$

for $2^{-j_3}(k_3 + 1) \leq \min(2^{-j_1}k_1, 2^{-j_2}k_2)$ with $K_a(j, k) = \exp(a(2^{-j}(k + 0.5)))$ and $H_j(a) = 1 - \exp(a2^{-j-1})$. The previous formulas can be easily extended to higher dimensions, and can be used to derive asymptotic behavior, e.g $|j_1 - j_2| \uparrow \infty$ and so on.

Another possible application of Proposition 1 is to non-stationary processes. A large variety of models, such as the bilinear model, autoregressive models with random coefficients, and the threshold model, have been proposed to take into account of the non-stationarity. To illustrate the use of cumulants, we restrict attention to piecewise stationary processes, i.e. a sum of independent stationary processes:

$$X(t) = \sum_{l=1}^{r} \mathcal{I}(u_l \leq t < u_{l+1})X^{(l)}(t), \text{ where } \mathcal{I}(A) = \begin{cases} 1, & \text{if } t \in A \\ 0, & \text{otherwise,} \end{cases}$$

and $X^{(l)}(t)$ are independent stationary processes and the change-points are equal to $-\infty = u_0 < u_1 < \cdots < u_r < u_{r+1} = \infty$. Because of linear properties of the cumulants, we have immediately that

$$CUM(X(t_1), ..., X(t_m)) = \sum_{l=1}^{r} \prod_{n=1}^{m} \mathcal{I}(u_l \leq t_n < u_{l+1})CUM(\underline{\mathbf{X}}^{(l)}(t)),$$

where $CUM(\underline{\mathbf{X}}^{(l)}(t)) = CUM(X^{(l)}(t_1), ..., X^{(l)}(t_m))$. Using Proposition 1, it follows that

$$CUM(\underline{\mathbf{W}}_{jk}) = \sum_{l=1}^{r} \int CUM(\underline{\mathbf{X}}^{(l)}(t)) \prod_{n=1}^{m} \mathcal{I}(u_l \leq t_n < u_{l+1}) \psi_{j_n, k_n}(t_n) dt_n$$

with $CUM(\underline{\mathbf{W}}_{jk}) = CUM(W_{j_1 k_1}, ..., W_{j_m k_m})$. The above expression shows that wavelet cumulants for piecewise stationary processes can be easily computed for compactly supported wavelets such as the Haar system, and when each process $X^{(l)}$ has simple cumulant functions (e.g. the CARMA process).

From the CAR example, we saw that wavelet cumulants are computable for the Haar system, but the resulting formula are not so easy to manipulate. Another approach is to use band-limited wavelets. Simpler expression of the wavelet cumulants can be derived. To illustrate this point, we look at the Meyer-type wavelet in the next proposition. In this case, the Meyer-type wavelet gives null or small wavelet cumulants within and across wavelet coefficient levels under simple conditions.

Proposition 2. *Let $\{X(t)\}$ be a stationary process that satisfies condition (1). Suppose that its m^{th}-order polyspectrum f_m is well defined, and the orthonormal basis $\{\psi_{jk}\}$ is generated by a mother Meyer-type wavelet. If there exists some integer j^* in $\{j_1, ..., j_m\}$ such that $\sum_{j \neq j^*} 2^j \leq 2^{j^*-2}$, then $CUM(W_{j_1 k_1}, ..., W_{j_m k_m}) = 0$. In addition, if $f_m(\omega)$ and $\tilde{\psi}(\omega)$ are both in C^p, then the cumulant at a fixed resolution level satisfies*

$$CUM(W_{j_1 k_1}, ..., W_{j_1 k_m}) = O(\max_r \prod_{s \neq r} |k_r - k_s|^{-p}).$$

Proposition 2 shows that the Meyer-type wavelet transform not only can remove the correlation inside the original signal, but in addition the higher-order cumulants are either null at distant scales or very small at a fixed scale. Walter's result [17] for the covariance is a special case of Proposition 2:

$$CUM(W_{j_1 k_1}, W_{j_2 k_2}) = \begin{cases} 0, & \text{for } |j_1 - j_2| > 1, \\ O(|k_1 - k_2|^{-p}), & \text{for } j_1 = j_2. \end{cases}$$

It is interesting to note that the results stated in Proposition 2 hold for any choice of F in the definition of the Meyer-type wavelet. An open problem is to determine if there exist some distributions F which will significantly reduce cumulants between wavelet coefficients. In this direction, Zayed and Walter [18] minimized the covariance between wavelet coefficients by using a bi-orthonormal wavelet basis that is a function of the original covariance.

In Propositions 1 and 2, different expressions of the wavelet cumulants were derived. A natural question is whether or not the cumulants of the original process can be expressed in terms of $CUM(W_{j_1 k_1}, ..., W_{j_m k_m})$. Thus, our next result is the converse of Proposition 1.

Proposition 3. *Suppose that $\{W_{jk}\}$ is a sequence of variables with all finite moments. If the process*

$$X(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} W_{jk}\psi_{jk}(t)$$

is well defined (in the mean-square sense), then we have

$$CUM(X(t_1), ..., X(t_m)) = \sum_{j_1,k_1=-\infty}^{\infty} \cdots \sum_{j_m,k_m=-\infty}^{\infty} \prod_{n=1}^{m} \psi_{j_n k_n}(t_n) CUM(\underline{\mathbf{W}}_{jk}).$$

§5. Conclusion and Future Work

In this paper, different relationships between wavelets and cumulants have been presented. Results show that the wavelet transform is not only a good tool to de-correlate a Gaussian process, but it also gives small higher-order cumulants of a non-Gaussian signal. The Meyer-type wavelet is particularly well-adapted for stationary processes since they give null wavelet cumulants at distant scales.

The combination of wavelets and cumulants has not yet been fully exploited. The statistical study of estimators of the bispectrum based on wavelets is of particular interest for application with real data sets. Also investigating the properties of non-linear and non-stationary times series models using bi-spectral methods and a wavelet decomposition approach needs further research.

Acknowledgments. This work was supported by NSF grants DMS-9815344 and DMS-9312686.

References

1. Brillinger, D. R., Some uses of cumulants in wavelet analysis, Nonparametric Statistics **6** (1996), 93–114.
2. Brockwell, P., Continuous-time ARMA processes, preprint.
3. Daubechies, I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
4. Donoho, D., S. Mallat and R. Von Sachs, Estimating covariances of locally stationary processes: rates of convergence of best basis methods, submitted to Annals of Statistics.
5. Ferreira, P. G., J. Magueijo and J. Silk, Cumulants as non gaussian Qualifiers, preprint.
6. Flandrin, P., On the spectrum of fractional Brownian motion, IEEE Trans. Info. Th., **35** (1989), 197–199.

7. Godfrey, M. D., An explorationary study of the bispectrum of economic time series, Appl. Statist. **14** (1965), 48–69.
8. Hasselman, K., W. Munk and G. MacDonald, Bispectra of Ocean Waves, in *Time Series Analysis*, M. Rosenblatt (ed.), John Wiley, New-York, 1963, 125–139.
9. Hinich, M. J., Testing for gaussianity and linearity of a stationary time series, J. Time Ser. Anal. **3** (1982), 169–178.
10. Johnstone, I. M. and B. W. Siverman, Wavelet threshold estimators for data with correlated noise, J. Roy. Statist. Soc. Ser. B **59** (1997), 319–351.
11. Lazear, G. D., Mixed-Phase wavelet estimation using fourth-order cumulants, Geophysics **58** (1993), 1042–1051.
12. Leporini, D., and J.C. Pesquet, High Order Properties of M-band Wavelet Packet Decompositions, IEEE Workshop on High-order Statistics, Banff, Canada, 1997.
13. Mallat, S., G. Papanicolaou and Z. Zhang, Adaptive covariance estimation of locally stationary processes, Annals of Statistics **26** (1998), 1–47.
14. Masry, E., The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion, IEEE Trans. Info. Th. **39** (1993), 260–264.
15. Meyer, Y., *Wavelets: Algorithms and Applications*, SIAM, Philadelphia, 1993.
16. Subba Rao, T. and M. Gabr, *An Introduction to Bispectral Analysis and Biliniear Time Series Models*, Lecture Notes in Statistics Vol. 24, Springer-Verlag, Berlin, 1984.
17. Walter, G. G., *Wavelets and Other Orthogonal Systems with Applications*, CRC Press, Boca Raton, 1994.
18. Zayed, A. I. and G. G. Walter, Characterization of analytic functions in terms of their wavelet coefficients, Complex Variables **29** (1996), 265–276.

Philippe Naveau and Doug Nychka
Geophysical Statistics Project
National Center for Atmospheric Research
B.P. Box 3000, Boulder CO 80307, USA
pnaveau@ucar.edu
nychka@ucar.edu

Peter Brockwell
Department of Statistics
Colorado State University
Ft. Collins, CO 80523-1877, USA
pjbrock@stat.colostate.edu